DOCUMENT RESUME

ED 185 064 TM 800 047

AUTHOR Haertel, Geneva D.; And Others
TITLE Social-Psychological Environments and Learning: A Quantitative Synthesis.
INSTITUTION Illinois Univ., Chicago. Chicago Circle Campus.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.; National Science Foundation, Washington, D.C.
PUB DATE Apr 79
GRANT NIE-G-78-0090; NSF-78-17374
NOTE 54p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, CA, April 8-12, 1979).

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Academic Achievement; *Classroom Environment; *Correlation; Elementary Secondary Education; Foreign Countries; *Outcomes of Education; *Predictive Measurement; Research Methodology; *Student Attitudes
IDENTIFIERS Australia; Canada; India; *Learning Environment Inventory (Anderson et al); United States

ABSTRACT

Student perceptions of their classroom environment correlated consistently with end-of-course cognitive, affective, and behavioral learning outcomes, with or without statistical controls for ability, pretests, or both measures. Twelve classroom observation studies, kindergarten to grade 12, that reported simple, partial, and part correlations between perceptions and outcomes were selected; 734 correlations were calculated from a combined total of 17,805 students in 823 classes in 4 countries. Cognitive measures included conventional achievement tests and tests of understanding, thinking and formal reasoning; affective criteria included interest, motivation, and self-concept tests; behavioral measures included self-report activity inventories and attendance. Most studies used the Learning Environment Inventory (LEI) to measure student perceptions. Learning outcomes were positively associated on the LEI with cohesiveness, satisfaction, task difficulty, formality, goal direction, democracy, and environment; and negatively associated with friction, cliqueness, apathy, disorganization and favoritism. Jackknifed regression equations showed that the magnitudes of the correlations depended on specific scales, level of aggregation, and nation; but not on sample size, subject matter, learning outcome, or statistical adjustments for ability and pretests. (Summaries of the 12 studies are appended). (Author/CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED185064

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Social-Psychological Environments and Learning: A Quantitative Synthesis

Geneva D. Haertel, Herbert J. Walberg, and
University of Illinois at Chicago Circle

Edward H. Haertel
University of Chicago

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Geneva Haertel

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Running head: Social-Psychological Environments

TM800 047

A paper presented at the 1979 Annual Meeting of the American Educational Research Association, San Francisco, California

Abstract

To estimate the sign and size of correlations of student perceptions of social-psychological environments of the classes with learning outcomes and adjusted gains, 734 correlations from a comprehensive collection of 12 studies of 10 data sets on 823 classes in eight subject areas containing 17,805 students in four nations were analyzed. A total of 31 of 36 hypotheses, theoretically-derived in 1969, are supported. Learning outcomes and gains are positively associated with Cohesiveness, Satisfaction, Task Difficulty, Formality, Goal Direction, Democracy, and the Material Environment and negatively associated with Friction, Cliqueness, Apathy, and Disorganization. Jacknifed regression equations show that the magnitudes of the correlations depend on specific scales, level of aggregation, and nation but not on sample size, subject matter, domain of learning outcome (cognitive, affective, or behavioral), or statistical adjustments for ability and pretests.

Social-Psychological Environments and Learning: A Quantitative Synthesis

During the past decade and a half, educational researchers and evaluators issued or published about one-hundred reports concerning student perceptions of the social-psychological dimensions of their classroom group such as cohesiveness, satisfaction, goal direction, difficulty, competitiveness, and friction. Reviews of this work (Randawa & Fu, 1973; Shulman & Tamir, 1973; Walberg, 1974, 1976, and Moos, 1979) discuss theoretical, methodological, and practical issues and conclude that such perceptions are useful as independent, mediating, and dependent variables in educational investigations in natural settings. Much of the research shows that social-psychological perceptual scales provide statistically-reliable sensitivity to educational treatments such as curriculum, teacher training, and instructional innovations as well as to project efforts to increase teamwork, cross-sex, cross-ethnic-group cooperation, and similar group properties. Other work reveals that such perceptions reflect and mediate teacher and student characteristics and that they provide diagnostically-valuable profiles of class and individual morale or climate.

The focus of the present work is the predictability of end-of-course cognitive, affective, and behavioral learning from mid-course social-psychological perceptions with and without statistical control for beginning-of-course measures or ability

both. Even if constructive perceptions of the social environment are considered worthy ends in their own right, it is important to determine if they are positively associated with learning gains and outcomes. Consistent, positive associations indicate that it is unlikely that learning is being traded off for more constructive social-psychological morale; under certain assumptions (Walberg, 1976), such associations may indicate causal connections between social-psychological perceptions and learning.

Although the reviews of research reported in twelve studies on ten independent, large-scale data sets show incremental, predictive validity of student perceptions, it has not seemed possible until recently to summarize the 734 correlations systematically and to answer such questions as: Which perceptions are most predictive? What learnings are most predictable? And, how does the predictability vary across such factors as grade levels of students, subject matter, and methodological characteristics of the studies? Jones and Fiske (1953), Light and Smith (1971), Gage (1978), Rosenthal (1976) describe a number of quantitative techniques,which are employed here, to synthesize the quantitative findings across studies and to provide answers to such questions (see Glass, 1978, for a critical exposition). As in quantitative summaries of empirical works in the natural sciences, the techniques are intended to provide estimates across investigations of: the consistency of observations or coefficents such as means, correlations,

and regression weights; their average magnitudes and margins of error; and their boundaries of application. The present application draws on the techniques developed by Glass (1978) for meta-analysis, by Mosteller and Tukey (1977) for obtaining appropriate error estimates when some of the data are not independent, and by ourselves for weighting independent data sources equally as well as estimating simultaneously the complete set of possible determinants of the correlation coefficients.

Method

Sample of Studies

A search was made of Dissertation Abstracts, Education Index, Psychological Abstracts, Social Science Citation Index, and, since much of the relevant research involved science curricula, the annual summaries of research sponsored by the National Association for Research in Science Teaching for the years 1963 through 1978. On-going, unpublished studies known by the authors or those cited in recent works were also considered for inclusion. All studies involving naturalistic classroom settings, kindergarten through twelfth grade, that report simple, partial, and part correlations between student perceptions of social-psychological climate of their classes and end-of-course learning were selected. The search and selection procedures yielded twelve investigations of ten data sets that report 734 correlations calculated from a combined total of 17,805 students in 823 classes (Table 1). The correlations

from three studies of a single data set (Walberg, 1969a and b,

---

Insert Table 1 about here.

---

and 1972) that explored predictability of learning across units of analysis and statistical-control techniques were counted as a single data source in the main regression analyses as explained in a subsequent section.

Features of Studies

Table 2 summarizes the key features of the twelve studies.

Insert Table 2 about here.

Information on each study includes a description of the environment measures used, the variables controlled, learning outcomes, and a brief statement of results.

In general the environment measures employed have internal-consistences reliabilities between .41 and .86. Regardless of the intended grade levels' of the instruments, the internal consistencies are approximately equal in magnitude.

The outcome measures include not only standardized achievement tasks, but affective and behavioral measures as well. Ten of the twelve included some cognitive measure of achievement, seven include affective or interest measures; and five employ a behavioral measure such as daily attendance.

Nine of the twelve studies are statistically controlled. Most studies control for the corresponding pretest; four studies control for student aptitudes such as IQ, and two control for instructional variables such as teacher attitude.

Although the focus of the present synthesis is the magnitude of the correlations for specific environment scales and learning outcomes, the multivariate results for sets of environment scales in the last column of Table 1 may be noted. Seven studies added sets of these scales to regression equations containing ability or pretest measures or both as controls, and reported the percentage increment in accountable variance. The average incremental variance accounted for on 19 learning outcomes is 20 percent with a range of 1 to 54 percent. Thus, regressions containing control and perceptual variables account for large amounts, in some cases, nearly all, of the either the total or reliable variance in learning outcomes.

Characteristics of Correlations

Information on eight characteristics was recorded for each simple, part, or partial correlation: national location of the study; grade level and number of students; unit of analysis; type of correlation; type of social-psychological perception; outcome domain; and content area of subject-matter (Table 3).

Insert Table 3 about here.

The continuous variables, grade level and number of students were grouped into class intervals to calculate frequencies and one-way analyses of variance, but were left in their full continuous precision for the regression analyses. The nominal variables such as location and unit of analysis were treated as categorical in the analyses of variance and converted to sets of binary (0.1) variables in the regressions, as explained in the discussion of the results.

Most of the variables listed in Table 3 are self-explanatory. However, a few deserve explanation. Unit of analysis refers to the level of aggregation used in the data analysis, i.e. student, sub-group, class, or school. With the exception of Moos and Moos (1978), and Bardsley (1976), all investigations employed the Learning Environment Inventory (LEI) in original or simplified or shortened form to reduce the 25-minute time required for the full 105 items for the 15 scales. For example, Perkins (1976) and Talmage and Walberg (1978) employed the 45-item My Class Inventory, an adaption of the LEI for elementary school. Since the five scales on the Moos and Moos instruments and adaptions of the LEI correspond closely, it was possible to code all scales to correspond to the LEI scales.

Outcome domain refers to the type of criterion measure used. Criterion learning measures are coded as either cognitive, attitudinal or behavioral. Types of cognitive measures include conventional multiple-choice achievement tests and tests of

understanding, critical thinking, and tests of formal reasoning. Attitudinal criteria include intruments such as interest measures and motivation and self-concept tests. Behavioral criterion measures include self-report activity inventories and absence rates.

The last item coded for each correlation coefficient is the outcome content area. Each correlation is coded for one of the following subjects: general science, life science, physical sciences, mathematics, social sciences, humanities, general achievement or miscellaneous subject areas which are aggregations across several of these categories.

Data Analysis

In the analyses, the magnitude of the correlation of environment scale and learning outcome is the dependent variable. The specific LEI scale, outcome domain, the sample size, unit of analysis, and the other explanatory factors are the independent variables. Several analytical techniques were used, beginning with a tabulation of the signs of the correlation by expected direction, proceeding through correlation-weighted one-way analyses of variance and culminating in series of study-weighted multiple regression analyses. The rationale of these procedures and details of their execution are explained in the next section but the weighting issue deserves discussion here.

The 734 correlation coefficients are by no means statistically independent since they arise from only twelve studies and ten independent data sets. Table 1 shows, moreover, that the number of correlations reported in the studies varies from 5 to 240. Weighting each correlation equally would give 48 times more weight to the larger study. For these reasons, special procedures were developed for the regressions to give each data set equal weight. "Jacknifed" estimates (Mosteller & Tukey, 1977; and Glass, 1978) of the regression weights were employed, to afford statistical independence for significance tests.

Results and Discussion

Directional Hypotheses

The 15 LEI subscales include some positive and some negative characteristics of the classroom environment. The positive subscales are: Cohesiveness, Satisfaction, Task Difficulty, Formality, Goal Direction, Democracy, Environment, and Competition. The negative scales are: Friction, Cliqueness, Speed, Apathy, Favoritism, Disorganization, and Diversity. In subsequent analyses the signs of correlations involving negative aspects of the classroom environment are reversed. Thus, the expected signs of correlations with all LEI scales, as coded for these analyses, are positive.

From social-psychological research, Walberg (1969b) derived 36 hypotheses concerning the direction of relations between

selected LEI scales and learning criteria, namely that Cohesiveness, Satisfaction, Task Difficulty, Goal Direction, Democracy, Diversity, and Environment would be positively correlated, and that Friction, Cliqueness, Apathy, Favoritism, Disorganization negatively correlated with the cognitive, affective, and behavioral learning. Data assembled for the present research synthesis permits a test of these hypotheses, by tabulating the number of signs across studies. The prediction for each of the 12 scales gave rise to 3 testable assertions, namely that correlations between that scale and measures in each of the three outcome domains would have the predicted sign. The resulting 36 testable assertions were evaluated as follows. Specifically, each study employing a given combination of scale and outcome domain was examined to determine whether the preponderance of coefficients is positive, negative, or evenly split. The three Project Physics studies that explored the consistency of correlations across class and individual units of analysis and analytic techniques were combined since they arose from a single data set (Tables 1 and 2). Ties were broken by assigning even splits the values plus, then minus, then plus, and so on. A tabulation of the results show that 31 of the 36 or 86 percent of the signs support the hypotheses; and the binonimal probability of an even split in a sign test is less than .001. Three of the five disconfirmations concern the Diversity subscale, which shows negative relation with out-

comes in all three domains, rather than the hypothesized positive relation.

Unweighted, Univariate Analyses

Table 3 reports the means, standard deviations, and frequencies of the 734 correlations grouped on the eight factors individually. These preliminary statistics and the F-tests for their corresponding correlation-weighted, one-way analyses of variance are intended to show trends, variations, and frequencies of the correlations for each of the eight factors separately, before presenting the results of the regressions. The latter provide more complicated, weighted and jacknifed estimates that weight each data set equally, employ a valid number of independent degrees of freedom, and combine all eight factors in a simultaneous analysis that controls each factor for the others and points to a parsimonious number of strong trends that summarize much of the significant variation in the magnitudes of the correlations.

The F-ratio for location of study is highly significant ($F = 35.58$, $df = 3,730$, $p < .001$). Table 3 shows that the correlations from studies in India and Canada are higher than those in the United States and Australia. This finding may be explained by reference to the specific studies from which these correlations were taken. The single study in India (Walberg, Singh, and Rasher, 1977) used only extreme groups of students, nominated as most and least studious in their classes.

The single Canadian study (Walberg and Anderson, 1972) is distinctive in that class, rather than student, was the unit of analysis. Aggregation effects may have raised correlations from these countries since analyses of collectivities usually yield stronger correlations.

The F-ratio comparing these correlations by grade level is not significant ($F = 0.10$, $df = 2,731$), which may be due in part to the unequal group sizes which resulted from the collapsing of the twelve grades into only three levels (elementary, junior high, high school). In the regression analyses discussed below, the grade levels were not collapsed; and grade level proved to be a strong predictor of correlation size.

Grouping the correlations by sample size yields a significant F-ratio ($F = 23.35$, $df = 3,730$, $p < .001$). The absence of linear trend across the different sample size categories, however, may suggest that the apparent effect of sample size is attributable to variations across studies in unit of analysis, type of correlation coefficient, type of outcome measure used, or other factors. The regression analyses were employed to test these possibilities.

With the exception of the mean correlation for "subgroups," the correlations show stronger relationships with larger units of analysis. Differences among the mean correlations with different units of analysis are clearly significant ($F = 19.44$, $df = 3,730$, $p < .001$). The anomolous value for subgroups may again be explained by reference to the peculiarities of the single

study using this unit of analysis by Walberg, Singh, and Rasher (1977). The method used to select a sample in this study may have given rise to unusally high correlation coefficients. Aside from the "subgroups" anomaly, the strength of the environment-outcome relation uncontrolled for the other seven factors, increases as larger and larger units are examined.

No significant differences were found between simple, part, and partial correlation coefficients ($F = .83$, $df = 2,731$). In some cases, partialling out ability or pretest scores or both and analyzing adjusted scores may increase precision and raise the correlation. In other cases, this increase may be more than offset by the attenuation due to the lowered reliability of the adjusted scores.

Differences among aspects of the environment measured, are highly significant ($F = 4.43$, $df = 14,719$, $p < .001$). The subscales which show the strongest relations to learner outcomes are Cohesiveness, Friction, and Satisfaction, all of which show average correlations of over .22 with outcomes. These are followed by Cliqueness, Difficulty, Apathy, Favoritism, Direction, Democracy, Disorganization, and Environment, with averages in the range from .16 to .12. The remaining four subscales, Speed, Formality, Diversity, and Competition, all show average correlations of less than .07.

The analysis of variance for outcome domain showed significant differences among the three outcome domains ($F = 19.67$, $df = 2,731$, $p < .001$). Higher correlations were observed with

outcomes in the cognitive domain than with those in either the attitudinal or behavioral domains.

The last of the eight one-way analyses of variance contrasts the eight content areas in which outcomes were related to environments ($F = 12.21$, $df = 7,726$, $p < .001$). Table 3 shows that the content areas in which outcomes are most predictable from environments are mathematics and the social sciences, followed by general science, the physical sciences, and the humanities. The category "general achievement" includes standardized test scores summed over several content areas. These indexes are all in the cognitive domain, and the relatively high mean correlation for this area may reflect primarily the exceptional reliability of such measures.

Regression Analyses and Jacknifed Estimates

Before the regression analyses, the categorical variables were replaced by sets of binary variables. Location, for example, was recoded into four variables, called "USA", "Canada", Australia", and "India". Each of these variables were given identifying values of either zero or one. For a correlation computed on a sample from the United States, "USA" is 1, and "Canada", "Australia", and "India" are each 0. "Australia" would be 1 and the rest 0, and so on. If the values of three of the variables are known, the fourth is also known and redundant; therefore only three need be entered in the regression, and the convention was adopted to omit the last value of each categorical variable in Table 4. The continuous variables, grade and sample size,

were left in their full metric precision.

In additon to recoding categorical variables, weights were introduced to equalize the contributions of the ten data sets to the estimates. Each correlation was given a weight proportional to the inverse of the number of correlations from its study. These weights were scaled so that the average weight for each cofficient was 1.00; thus the sum of the weights is the number of coefficients in the sample.

After the recoding and dropping of one variable from each binary set, 33 variables were available for the regression analysis: 3 variables for location, 1 variable for grade level (not recoded), 1 variable for sample size (not recoded), 3 variables for unit of analysis, 2 variables for correlation type, 14 variables for subscales, 2 variables for outcome domain, and 7 variables for content area. Consideration of variables representing interactions of LEI scales with sample and study characteristics introduce still more variables. It would be quite unusual if all of these variables possessed significant power to predict the size of the correlation coefficients after controlling for all the other variables. The problem, therefore, was to decide which of the 33 variables and additional interactions should be included in the final regression equation. To screen the variables, a multi-stage procedure was employed, whereby weak or colinear predictors were successively eliminated. First, a run was made using all 33 predictors. The those with F-ratios of less than 1.00 were eliminated, and a second run was made. Then all remaining variables with

F-ratios less than 2 were dropped, and the remaining variables were used in a third run. Variables in this run which showed F-ratios of less than 4 were eliminated, and a fourth run was conducted. This run included only 18 of the original 33 variables. At this point, sets of product variables were introduced that measure the influence of interactions of significant LEI scales with grade level and unit of analysis. To determine which of these product variables possessed additional explanatory value, a stepwise procedure was employed, in which the 18 variables already identified were forced into the equation and product terms were then entered one at a time, using an F ratio of 4.00 as the criterion for entering each new variable. The final criterion corresponded closely to significance at the .05 level. A total of 32 variables were included in the final equation. In addition to the 18 variables in the last reduced equation, 14 cross-product terms, representing interactions, were introduced.

Conventional significance tests computed for the regression coefficients assume that the correlations are statistically independent. Since in all cases several or many correlations are taken from the same study, this assumption is invalid. Accordingly, the significance of the coefficients was estimated using the "jacknife" (Mosteller & Tukey, 1977) that treats studies as independent, but makes no assumptions about the independence of multiple measures of association taken from the same study. To apply the jacknife, the final regression

equations for the reduced and cross-product models were computed ten times, each time omitting one of the data sets. These ten regression equations, together with the original equation were then used to estimate the metric b-weights, and their standard errors. Jacknifing yielded an independent estimate of each b-weight, as well as a corresponding t-ratio on nine degrees of freedom.

Reduced Model

As shown in Table 4, jacknifing reveals 10 of the 18 coefficients from the reduced equation and 15 of the 32 coefficients product model significant at the .05 level. The jacknifed estimates are similar to the original regression estimates; but the t values are somewhat lower on average, which indicates some inflation of significance levels occurs as a result of the non-independence of correlations from the same study when analyzed by conventional methods. The majority of the independent variables in the jacknifed reduced equation, however, are significnatly weighted in predicting the magnitude of the correlations between LEI scales and learning outcomes.

Among the variables representing location of the study, USA had a b-weight significantly less than zero in the context of the jacknifed reduced model. Holding other things constant, correlations taken from studies conducted in the United States averaged .22 less than those from studies aboard.

Grade level shows a small but persistant posiitve relationship to correlation size; the average correlation rises roughly .05 per year. Older high school students have longer

school experience, usually attend more classes each week, and may thus be more astute raters of the class learning environment.

With other variables in the basic equation held constant, a clear trend appears in the size of correlations as a function of unit of analysis. Although only the variable representing class as unit-of-analysis is significant by the jacknife procedure, t-tests for student and subgroups as unit are nearly significant and reveal, in the context of the other variables, a monotonic increase in the magnitude of the correlations with increasing aggregation from student to subgroup to class to school as unit analysis.

The strength of the LEI scale-outcome relation is significantly higher for seven of the scales than the others. The absolute correlations of learning with Cohesiveness, Friction, Satisfaction, Favoritism, Goal Direction, Democracy, and Environment average .30 to .38 higher than those for the other eight scales in the context of the regression.

It is notable that type of correlation is not significant in any of the variance or regression analyses: the differences among simple correlations of the perceptual scales and learning outcomes and part and partial correlations with ability or cognitive, affective, or pretests both types of measures controlled are statistically undetectable. The unreliability of adjusted scores may compensate for increased control for aptitudes. However, since Anderson and Walberg (1974) show that IQ contributes little to the prediction of regression-adjusted gains in cognitive, affective, and behavioral learning

in several data sets whereas LEI scales contribute substantially, the scales appear valid in tapping the mediating or causal social-psychological forces in the determination of learning that are independent of aptitudes and pretests.

It is also notable that learning domain is not significant: the correlations of perceptions and cognitive outcomes do not differ significantly from those involving affective and behavioral learning outcomes. Thus, it appears that constructive aspects of class morale are equally associated with outcomes in all three domains rather than being associated with benefits in one domain sacrificed for losses in another.

The reduced model in Table 4 may be used to estimate the sizes of correlations to be expected under specific conditions in future research. For variables that do not significantly enter the equation--notably, schools as units and eight LEI scales, the scale - learning estimate is the constant, .42 at the Kindergarten level. The b-weights for the binary predictors can be added to the constant for other specific estimates; for example, the estimate for the correlation of Cohesiveness and learning in Canada and India with classes as units is .42 + .36 + (-46) = .32 at the Kindergarten level; the same estimate for the United States must be reduced by .22 which equals .10. For the corresponding U.S. estimate at the twelfth grade, 12(.05) can be added to .10 which equals .70. Least confidence can be placed in the estimates of correlations for elementary grades

since only two of the ten data sets were obtained at this level (Tables 2 and 3). Extreme caution must be taken in extrapolating estimates beyond the data ranges given in Tables 2 and 3; caution is also required for interpolated estimates within sparsly sampled data regions such as those for India, elementary and junior-high grades, and part correlations (Table 3) as well as sub-spaces of these regions, which deserve additional empirical investigation to enlarge the areas of confident estimation.

Product Model

Entering significant interactions of grade level and unit of analysis with LEI scales changes the magnitude and significance of the b-weights but not their sign (Table 4). For example, although the estimated b-weight for grade level is about the same, it is not significant. The negative signs of the significant products of grade level with Cohesiveness and Satisfaction (footnote to Table 4) indicate that the non-significant rise in perception-learning correlations with grade level is significantly reversed in the cases of these two affective perceptions of class morale. Thus, correlations of these affective aspects of the social environment appear to rise less steeply in their relation to learning across the grade levels as compared to other aspects such as Goal Direction, Democracy, and Task Difficulty. Confirmations of these trends in additional empirical investigations particularly those in the elementary grades would support the interpretation that organizational and task aspects of the

social environment strenthen and affective aspects weaken relative to one another with increasing grade level.

Six interactions of unit of analysis and LEI scale are significant (Table 4 footnote). As in the case of grade level, they call for qualifications of the general relations of perceptions and learning; these interactions reveal stronger associations of specific scales than others at certain levels of analysis. Since only one study used schools as units and another used sub-groups within classes, these interactions do not warrant much interpretation until further empirical studies are conducted. Cronbach (1976 and personal communications) and Walberg (1976) have discussed substantive and methodological issues of varying empirical relations across units of analysis. Future investigations can contribute to the understanding of these complexities by conducting parallel analyses using the student, sub-group, class, and school, as the units of analysis.

Although the individual b-weights for the product terms should be interpreted cautiously, on the whole, inclusion of statistically significant product terms improved the fit of the model. This is indicated in Table 4 by the increase in the multiple correlation from .56 for the reduced model to .71 for the product model. Illustrative coefficient estimates derived from each of the two models are presented in Table 5, together with corresponding observed coefficients taken from two of the studies. These are observed and estimated coefficients for U.S. classes, computed for grades 4 (elementary) and 11 (high school). The

signs and magnitudes of the observed and estimated coefficients in Table 5 are in good agreement, given the standard errors for the two models as reported in Table 4. It should be noted that the estimates for high school samples are clearly more accurate, and more empirical work on elementary samples is in order. As would be expected, the estimates derived using the product model are in somewhat closer agreement with the observed values than those derived from the reduced model. The difference in goodness of fit of the two models is not large, however, and for most purposes either set of estimates provides a reasonable summary of the data structure as well as expected sizes of correlations for future empirical investigations.

Conclusions

Across ten data sets from four countries and in a variety of samples, subject matters, and methodological approaches, perceptual aspects of the social-psychological environment of learning are correlated consistently in sign in their relation to cognitive, affective, and behavioral learning outcomes with or without statistical controls for ability, pretests, or both measures. Specifically, these learning of outcomes are positively associated with Cohesiveness, Satisfaction, Task Difficulty, Formality, Goal Direction, Democracy, and Environment and negatively associated with Friction, Cliqueness, Apathy, Disorganization and Favoritism. As a set, these perceptions contribute substantially accounting for variance in learning outcomes beyond the variance accounted for by ability and pretest measures.

The correlations differ significantly in magnitude across perceptual scales, units of analysis, nations, and grade levels, and combinations of scales and the other factors. Although these differences deserve further empirical investigation, the theoretical plausibility and incremental predictive validity of the scales as well as their utility for further research and evaluation seem warranted. Their causal relation to learning is plausible but unproven. Educators who doubt the causal relation, however, or who believe in the inherent value of learning environment properties as ends in themselves rather than as means to standard outcome measures may not need to fear sacrificing one for the other since they appear to go together.

Footnotes

1. The authors thank Maurice J. Eash and Harriet Talmage for institutional support. The research presented in this article was supported by the National Institute of Education (Grant No. NIE-G-78-0090) and the National Science Foundation (Grant No. NSF-78-17374); the points of view and opinions stated do not necessarily represent the official position or policy of either agency.

2. The authors thank Professor Chad Ellet at the University of Georgia for providing data from an unpublished dissertation (Perkins, 1976).

References

Bardsley, W. N. Student alienation and commitment: A multivariate analysis of the effects of home and school environments. Unpublished doctoral dissertation, Australian National University, 1976.

Cronbach, L. J. Research on classrooms and schools: Formulation of questions, design, and analysis. Occasional paper, Stanford Evaluation Consortium, Stanford University, California, 1976.

Fraser, B. Evaluation of a science-based curriculum. In H. J. Walberg (Ed.), Educational environments and effects. Series on Contemporary Educational Issues. Chicago: University of Chicago Press, 1979.

Gage, N. L. The scientific basis of the art of teaching. New York: Teachers College Press, 1978.

Glass, G. V. Integrating findings: The meta-analyses of research. In L. S. Shulman (Ed.), Review of Research in Education, Vol. 5. Itasca: Peacock, 1978.

Jones, L. V. and Fiske, D. Models for testing the significance of combined results. Psychological Bulletin, 50, (3), 1953, 375-382.

Light, R. J. and Smith, P. V. Accumulating evidence: Procedures for resolving contradictions among different research studies. Harvard Educational Review, 1971, 41, 421-471.

Moos, R. H. Educational climates. In H. J. Walberg (Ed.), Educational environments and effects. Series on Contemporary Educational Issues. Chicago: University of Chicago Press. 1979.

Moos, R. and Moos, B. Classroom social climate and student absences and grades. Journal of Educational Psychology, (in press).

Mosteller, F. and Tukey, J. W. Data analysis and regression: A second course in statistics. Reading: Addison-Wesley, 1977.

Perkins, M. L. A canonical correlational analysis of the relationships among school climate, teacher morale, and educationally-relevant performance of fourth grade students. Unpublished doctoral dissertation. University of Georgia, 1976.

Randhawa, B. S. and Fu, L. L. W. Assessment and effect of some classroom environment variables. American Educational Research Journal, 1973, 43, 303-321.

Rosenthal, R. Experimenter effects in behaviroal research. New York: Irvington Publishers, 1976.

Shulman, L. S. and Tamir, P. Research on teaching in the natural sciences. In R. M. W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand Mc-Nally, 1973.

Talmage, H. and Walberg, H. J. Naturalistic decision-oriented evaluation of a district reading program. Journal of Reading Behavior, 1978, X, 185-195.

Tisher, R. P. and Power, C. N. The effects of classroom activities, pupil's perceptions and educational values in lessons where self-paced curricula are used. Research Report. Australian Advisory Committee on Research and Development in Education, 1975.

Walberg, H. J. Social environment as a mediator of classroom learning. Journal of Educational Psychology, 1969, 60, 443-448. (a)

Walberg, H. J. Predicting class learning: An approach to the class as a social system. American Educational Research Journal, 1969, 6, 529-542. (b).

Walberg, H. J. Social environment and individual learning: A test of the Bloom model. Journal of Educational Psychology, 1972, 63 (1), 69-73.

Walberg, H. J. (Ed.) Evaluating educational performance: A sourcebook of methods, instruments, and examples, Berkeley: McCutchan, 1974.

Walberg H. J. Psychology of learning environments. In L. S. Shulman (Ed.), Review of research in education, Vol. 4. Itasca, Illinois: F. E. Peacock, 1976.

Walberg, H. J. and Anderson, G. J. Classroom climate and individual learning. Journal of Educational Psychology, 1968, 59 (6), 414-419.

Walberg, H. J. and Anderson, G. J. Properties of the achieving urban class. Journal of Educational Psychology, 1972, 63, (4), 381-385.

Walberg, H. J. and Anderson, G. J. Learning environments. In H. J. Walberg (Ed.) Evaluating educational performance: A source book of methods, instruments, and examples. Berkeley, California: McCutchan, 1974.

Walberg, H. J., Singh, R., and Rasher, S. P. Predictive validity of student perception: A cross-cultural replication. American Educational Research Journal, 1977, 14, (1), 45-49.

Table 1

Statistics on Twelve Studies

Study	Number of Correlations	Number of Classes	Number of Students
Bardsley (1976)	7	30	374*
Fraser (1979)	11	153*	541
Moos and Moos (1978)	10	19*	375
Perkins (1976)	5	108	3700
Talmage and Walberg (1978)	5	59*	1600
Tisher and Power (1975)	240	20	315*
Walberg and Anderson (1968)	22	76	2600*
Walberg and Anderson (1972)	150	64*	1600
Walberg, Singh and Rasher (1977)	60	150*	3000
Subtotal	510	679	14,105
Walberg (1969a)	84	144*	3700
Walberg (1969b)	84	144*	3700
Walberg (1972)	56	144	3700*
Subtotal	224	144	3700
Total	734	823	17,805

Note: Since three studies (Walberg, 1969 a & b, 1972) analyzed a single data set to explore predictability across units of analysis and other methodological variations; the correlations from these studies were combined in the regression analyses to give each data set equal weight.

Table 1 continued

*The units of analysis for each study are indicated with an asterisks. See text on the sub-group analysis in the Walberg, Singh, and Rasher (1977) study. See text on the use of school as unit of analysis in the Perkins (1976) study.

Table 2

Design Characteristics and Results of Twelve Studies

Author (date)		Environment Measures	Outcome Measures	Controls	Results
Bardsley (1976)	374 senior high students in Australia. Student is unit of analysis.	A single questionnaire item which corresponds to the Formality subscales of the Learning Environment Inventory (LEI).	Series of adjustment variables: misfeasance, self-estrangements, social powerlessness, value isolation, meaninglessness, social isolation, task powerlessness.	None	Correlation of "Rules" item with subscales are weak: misfeasance = 01; self-estrangement = .14, social powerlessness = .24, value isolation = .00, meaningfulness = .06, social isolation = .04 task powerlessness = -.07
Fraser (1978)	41 students in 20 seventh-grade general science classes, Melbourne, Australia, 10 experimental and 10 traditional curriculum; 153 sub-groups classified by sex, socio-economic status, and ability as within classes units	Modified, 55-item of the LEI containing 9 scales with internal consistencies from .50 to .80	Seven cognitive and affective outcome measures ranging in internal consistency from .63 to .91.	Pretests, student attitudes, and instruction	In guided stepwise regression, aptitude alone produced Rs of from .48 to .76; instruction raised these by .16 to .26; and LEI raised R additionally from .17 to .47. Total R ranged from .81 to .86.

Table 2 continued

Moos and Moos (1978)	19 representative classes in a U.S high school; classes as units	90-item, 9-scale Classroom Environment Scale with internal consistencies from .67 to .86	Attendance and average grade given in class	None	Absences correlated with Competition and Teacher Control; grades correlated positively with Involvement, Affiliation, and Teacher Support and negatively with Rule Clarity and Teacher Control.
Perkins (1976)	About 3,700 elementary school students in grade 4 from 42 U.S schools; School was the unit of analysis	My School which is identical to My Class contains 5 subscales: Cohesiveness, Competitiveness, Friction, Difficulty and Satisfaction.	Five subscales of the Iowa Test of Basic Skills: Vocabulary, Reading, Language Skills, Work Study Skills, Math Skills; and average daily attendance	Teacher attitude variance partialed out.	Positive relationship between performance on My Class and student performance on achievement tests when teachers' perception of the school environment is removed.
Talmage and Walberg (1978)	About 1600 elementary school students in grades 1, 2, 3, and 6; 59 classes as units	My Class, an elementary school version of the LEI with 5 scales as specified above, reliabilities from .54 to .77.	Science Research Associates Reading Test Total Scores	SRA alternate form pretest given one year earlier	Pretest-posttest r. of .87 raised to R of .93.

Table 2 continued

Tisher and Power (1975)	315 junior high students in grade 9 from 20 classes in Australia. Student was the unit of analysis	Modified version of 15 LEI scales with internal consistencies from .53 .82.	Achievement in content areas, attitudinal measures, and satisfaction with teaching methods.	7 selected pretests: 1)ecology achievement 2)population ecology 3)pollution 4)population 5)resources 6)interest 7)learning	Attitudes: Pre-test posttest r of .33 raised to .48, satisfaction (3 scales): .32 raised to .52, .20 raised to .47, .22 raised to .46
Walberg and Anderson (1968)	2,100 students in 76 high school physics classes in U.S; student as unit	80-item, 18 factor analytically-derived scales on the Classroom Climate Questionnaire with internal consistencies ranging from .41 to .86.	9 cognitive, affective, and behaviroal posttests regression-adjusted for corresponding pretests ranging in internal consistency from .61 to .86.	None	20 percent of the 162 intercorrelations significant at .05; achievement positively correlated with Intimacy, negatively correlated with Goal Diversity and Social Heterogeneity; understanding negatively correlated with Stratified; affect positively correlated with Democratic, Goal Direction, Formal, Egalitarian, and Satisfaction and negatively with Stratified, Friction, Disorganized, Social and Interest Heterogeneity.

Table 2 continued

Walberg and Anderson (1972)	About 1600 students in 64 Montreal high school classes in 8 subjects; class as unit.	The fifteen LEI scales. Reliabilities from .58 to .86.	Standardized Quebec High School Learning Examinations; internal consistencies range from .70 to .80.	IQ	In split-sample double cross-validations, r with controls raised to R with LEI and cross-validated: .42 to .87 and .67 in sample A and .24 to .78 and .43 in sample B.
Walberg, Singh, and Rasher (1977)	Somewhat less than 3,000 students in 300 studious and non-studious sub-groups of 10 students each in 83 science and 67 social studies classes in Rajasthan, India High Secondary schools; sub-group as unit.	The fifteen LEI scales. Reliabilities from .58 to .86.	100-item multiple-choice achievement tests geared to standard curriculum; general science internal consistency, .67; social studies, .81.	IQ	In general science, r of .63 raised to R. of .82; in social studies, .61 to .90.

Table 2 continued

Walberg (1969a)	3,700 students in national sample of 144 high school physics classes; class as unit	Learning Environment Inventory: Fourteen scales with internal consistencies from .58 to .86 for individuals and .43 to .84 for classes	Test on Understanding Science, internal consistences, .76; Welch Science Process Inventory, Physics Achievement Test, .77; Academic Interest Measure, .91; Pupil Activity Inventory, .80; Semantic Differential, .86; all given at the beginning and end of a one-year course.	IQ, pretest achievement and interest	On three cognitive criteria, median R with controls of .66 raised to .72; on three non-cognitive criteria, median R raised from .40 to .51; on achievement, .71 to .73.
Walberg (1972)	Same as Walberg (1969a), but student as unit	Same as above	Selected pretests and posttests on understanding, achievement, interest, and activities.	Corresponding pretest	R with controls raised: from .68 to .69 for understanding, .73 to .75 for achievement, .70 to .70 (not significant) for physics interest, and .75 to .76 for voluntary physics activities

Table 2 continued

Walberg (1969)	Same as Walberg (1969a), but class as unit.	Same as above	Same as above	Corresponding pretest	R of regression-residualized gain score with 14 LEI scales: .45 for cognitive criteria; .41 for non-cognitive criteria; .43 for achievement.

DOC 42

Table 3

Descriptive Statistics for Correlations

Between Educational Outcomes and Learning Environments

Factor	Mean Correlation	Standard Deviation	Frequency
Location*			
USA	.10	.15	266
Canada	.26	.37	150
Australia	.06	.09	258
India	.32	.49	60
Grade Level			
Elementary	.12	.36	10
Junior High	.11	.13	11
High School	.14	.26	713
Sample Size*			
40-299	.25	.40	120
300-499	.06	.12	257
500-999	.05	.08	67
1,000-3703	.18	.27	290
Unit of Analysis*			
Students	.07	.10	325
Subgroups	.29	.46	71
Classes	.17	.29	333
Schools	.30	.31	5

Table 3 continued

Type of Correlation			
Simple	.15	.29	442
Part	.11	.13	11
Partial	.12	.20	281
Learning Environment Scale*			
Cohesiveness	.23	.27	50
Friction[a]	.23	.23	53
Cliqueness[a]	.12	.19	46
Satisfaction	.22	.21	54
Speed[a]	.02	.31	48
Task Difficulty	.13	.24	50
Apathy[a]	.14	.32	48
Favoritism[a]	.16	.16	46
Formality	.06	.26	57
Goal Direction	.17	.23	51
Democracy	.17	.24	50
Disorganization[a]	.13	.22	50
Diversity	.02	.73	47
Environment	.18	.26	49
Competition	.06	.38	35
Outcome Domain*			
Cognitive	.17	.33	403
Attitudinal	.10	.11	284
Behavioral	.07	.13	47

Table 3 continued

Content Area*			
General Science	.12	.25	133
Life Sciences	.05	.11	165
Physical Sciences	.12	.19	279
Mathematics	.38	.37	15
Social Sciences	.34	.50	60
Humanities	.15	.35	35
General Achievement	.25	.28	40
Miscellaneous	.08	.08	7

*Factors significant at the .001 level are indicated with an asterick.

[a]Signs of correlations reversed for these LEI scales.

Table 4
Conventional and Jacknifed Regression Statistics for the Two Models

Variable	Reduced Model				Cross-Product Model			
	Conventional		Jacknifed		Conventional		Jacknifed	
	b-weight	t	b-weight	t	b-weight	t	b-weight	t
USA	-.20	-6.94	-.22	-2.51	-.23	-8.72	-.25	-2.33
Australia	-.16	-4.34	-.12	-1.05	-.18	-5.26	-.19	-1.54
Grade Level	.03	5.41	.05	2.80	.06	8.42	.06	1.01
Student as Unit	-.37	-6.36	-.82	-2.22	-.18	-3.03	-.28	-1.61
Subgroup as Unit	-.30	-5.72	-.76	-2.09	-.21	-3.94	-.42	-1.98
Classroom as Unit	-.39	-8.95	-.46	-4.57	-.26	-5.41	-.27	-3.09
Cohesiveness	.37	9.51	.36	3.23	.91	9.76	.85	16.38
Friction[a]	.37	9.40	.37	2.64	.62	6.40	.42	1.19
Cliqueness[a]	.20	3.26	.17	1.72	.12	2.18	.08	.77
Satisfaction	.36	9.92	.38	4.00	.82	9.50	.61	9.10
Task Difficulty	.27	6.80	.27	2.11	.72	10.12	1.03	3.36
Apathy[a]	.11	2.33	.11	.96	.16	3.51	.14	1.62
Favoritism[a]	.23	3.86	.21	2.46	.15	2.86	.12	1.24
Formality	.16	4.24	.12	1.39	.01	.20	.04	.60
Goal Direction	.30	6.60	.34	3.05	.15	3.10	.20	2.32

Table 4 continued

Variable	b-weight	t	b-weight	t	b-weight	t	b-weight	t
Democracy	.30	5.93	.32	3.04	.14	2.79	.17	2.18
Disorganization[a]	.17	3.73	.16	1.63	.08	1.89	.06	.74
Environment	2.8	6.14	.31	3.06	.13	2.16	.12	1.37
Constant	.11		.42	1.19	-.20		.05	.18
Standard Error	.247		.259		.214		.225	
R	.571		.533		.709		.691	

[a]Signs of correlations for these LEI scales reversed in all analyses.

Note: For the jacknifed estimates, t-values of 2.26, 3.25, and 4.78 are significant at the .05, .01, and .001 levels, respectively. The jacknifed b-weights (and t-values) for cross-product terms are as follows: grade level by Cohesiveness -.03(3.53), Friction -.02(.64), Satisfaction -.05(3.78); Student as Unit of Analysis by Cohesiveness -.36(2.14), Task Difficulty -.98(2.99); Subgroup as Unit of Analysis by Task Difficulty -.30(3.20), Apathy -.99(12.26), Formality .39(6.66), Goal Direction .17(.84), Democracy .19(2.44), Environment .24(1.43); Class as Unit of Analysis by Cohesiveness -.33(5.95), Friction .41(1.75) and Task Difficulty -.94(3.01).

Table 5

Correlations of Learning Environment with Achievement at Two Grade Levels for U. S. Classes

	Elementary			High School			
	Reduced Estimate	Product Estimate	Observed	Reduced Estimate	Product Estimate	Observed History	Observed Physics
LEI Scale							
Cohesiveness	.27	.17	.00	.63	.38	.81	.93
Friction[a]	-.29	-.52	-.37	-.65	-.80	-.90	-.75
Cliqueness[a]				-.45	-.27	-.74	.04
Satisfaction	.30	.38	.11	.66	.45	.63	.63
Task Difficulty	.19	-.14	-.27	.55	.28	.26	.53
Apathy[a]				-.39	-.33	-.86	-.79
Favoritism[a]				-.50	-.31	-.59	-.03
Formality				.39	.23	-.12	.55
Goal Direction				.62	.39	.66	.44
Democracy				.59	.36	.61	.50
Disorganization[a]				-.44	-.25	-.71	-.42
Environment				.59	.55	.80	.88
All others	-.08	-.23	.51	.28	.19	-.15	-.57

Table 5 continued

[a]Signs of correlations for these LEI scales were reversed in all analyses. The original signs, however, have been restored in this table. Observed correlations were obtained from Talmage and Walberg (1978) and Walberg and Anderson (1972).